REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
11/10/06	Final	15/01/03 - 30/06/06

4. TITLE AND SUBTITLE
Detecting, Classifying, and Handling Contradictions in a Large, Dynamic Information Environment

5a. CONTRACT NUMBER

5b. GRANT NUMBER
F496200310065

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Donald Perlis
Michael Anderson
Darsana Josyula
Waiyian Chong
Scott Fults

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Maryland, College Park
College Park, MD 20742

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF, AFRL
AF Office of Scientific Research
875 N. Randolph Street, Room 3112
Arlington VA 22203
Dr Berklotz

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR/PKA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-SR-AR-TR-06-0448

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approve for Public Release: Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A new approach to perturbation tolerance was identified---the Meta-Cognitive Loop (MCL)---for responding to contradictions and other anomalies in complex settings. Further investigations with MCL included identifying architectural requirements, and applying MCL to various domains including reinforcement learning, commonsense reasoning, and a task-oriented natural-language interface system. A series of experiments empirically demonstrated the efficacy of MCL in improving the perturbation tolerance of certain machine learning techniques, including Q-learning, SARSA and Prioritized Sweeping. Formal metrics were given for measuring the complexity, dynamicity and overall difficulty of test domains, which allow for derivative measures of perturbation tolerance. A semantics was developed for Active Logic---the underlying logic on which MCL's contradiction handling is based---in the propositional case.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		11	Donald Perlis
					19b. TELEPHONE NUMBER *(Include area code)* (301) 405-2685

FINAL REPORT

Project: Detecting, classifying and handling contradictions in a large, dynamic information environment

PI: Donald Perlis and Michael Anderson (co-PI)
University of Maryland
College Park, MD 20742

Agreement #F496200310065

ABSTRACT:

A new approach to perturbation tolerance was identified---the Meta-Cognitive Loop (MCL)---for responding to contradictions and other anomalies in complex settings. Further investigations with MCL included identifying architectural requirements, and applying MCL to various domains including reinforcement learning, commonsense reasoning, and a task-oriented natural-language interface system. A series of experiments empirically demonstrated the efficacy of MCL in improving the perturbation tolerance of certain machine learning techniques, including Q-learning, SARSA and Prioritized Sweeping. Formal metrics were given for measuring the complexity, dynamicity and overall difficulty of test domains, which allow for derivative measures of perturbation tolerance. A semantics was developed for Active Logic---the underlying logic on which MCL's contradiction handling is based---in the propositional case.

OBJECTIVES:

(No change from original proposal) We will develop a rule-based reasoning system with the ability to detect, evaluate, and adjudicate contradictions. A typology of contradiction will be proposed; formal and implemented rules for addressing contradictions will be generated; and a complex testbed application will be implemented. The aim of the proposed research is to produce a reasoning agent which can recognize contradiction, repair it when possible, do this effectively for a wide range of common contradiction types, but continue to effectively reason even when repair efforts fail.

ACCOMPLISHMENTS:

1. Formal specification of a language for expressing the contents of the knowledge base (KB) of the general purpose Active Logic (AL) reasoner---Alma/Carne. We also developed a meta-theory that provides a formal specification of how the knowledge base evolves during reasoning. Different properties that the evolving knowledge base exhibits have been proven formally within the meta-theory.

2. Development of an Active Logic based meta-cognitive architecture--DIRECTOR--for implementing a universal interfacing agent that can connect multiple task-oriented systems to a user. DIRECTOR represents beliefs, desires, intentions, expectations, observations and achievements explicitly in order to reason about an agent's mental attitudes.

3. Development of a theory, ALFA, based on Active Logic, to reason about actions and mental attitudes. ALFA has the ability to reason about and execute concurrent actions, as well as non-concurrent actions requiring concurrent results. ALFA provides a meta-cognitive solution to dealing with the problems of interacting preconditions and opposing effects when concurrent actions are involved. In ALFA, a meta-cognitive process marks desires, intentions and expectations as achievable or not, based on real-time conditions. Achievable

desires become intentions, achievable intentions are acted upon, actions cause expectations and these expectations create further desires. ALFA provides an automatic and adaptive mechanism for intention reconsideration characterized by the following two properties: (i) only achievable intentions result in an actual action being initiated and (ii) an unachievable intention can result in a revised intention. A similar meta-cognitive process drops intentions that have been achieved or have become futile. ALFA essentially deals with two broad categories of contradictions: (i) those that get noted as direct contradictions in the underlying logic and (ii) those that get noted by the meta-cognitive process as not achievable.

4. Further refinement our specification of the perturbation-tolerant reasoning framework we call the Meta-Cognitive Loop (MCL). In our view, systems should have expectations attached to (generated for) every action they take, both locally (at the level of system components), and globally (for actions taken by the system), as well as more general expectations for the sorts of things that should happen when they are operating properly. This should allow for the detection of a wide range of perturbations.

5. A series of experiments comparing the performance of MCL-enhanced reinforcement learners with standard reinforcement learners in a changing world. The MCL-enhanced reinforcement learner, in contrast to the standard one, was able to notice that something was wrong, and take specific steps to overcome the problem. We found that the MCL-enhanced learners were able to recover much more quickly from large changes in the world, thereby maintaining much better overall performance.

6. Formalization of the metrics earlier developed for measuring the complexity of an environment (including such things as its degree of change over time, and degree of difference from place to place), and a demonstration of how these can be used to measure the perturbation-tolerance of an autonomous system.

7. Definition of a semantics for Active Logic (AL) in the propositional case. AL is different from classical logics in that it has special rules defining inference in temporal terms, and allows for controlled reasoning in the presence of contradictions. This requires appropriate modifications of the classical notions of a model, of logical consequence, and of the soundness of inference rules. Inspired by the notion that until an agent notices that a set of beliefs is contradictory, that set seems consistent (and the agent therefore reasons with it as if it were consistent), we introduced an ''apperception function'' that represents an agent's limited awareness of (the consequences of) its own beliefs, and serves to modify inconsistent belief sets so as to yield consistent sets. Using these ideas, we introduced a new definition of model and of logical consequence, as well as a new definition of soundness such that, when reasoning with consistent premises, all classically sound rules are sound for active logic. However, not everything that is classically sound remains sound in our sense, for by classical definitions, all rules with contradictory premises are vacuously sound, whereas in active logic not everything follows from a contradiction.

8. Improvements to the natural-language human-computer interface ALFRED, expanding the amount of self-monitoring it is able to perform, thereby increasing its ability to detect and recover from perturbations, including unknown words and violated expectations for event timing (such as user responses or domain actions). We have also enhanced and expanded the number of domains to which ALFRED can connect. These now include: toy trains, email, restaurant database, electronic home control, pool, home theater, furniture, draughts and chess. It can also interact with a universal house domain that includes all the nine domains as sub-components.

9. Design and implementation of a perturbation-tolerant, simulated khepera robot that is able to notice when navigational failures (such as collisions) take place, and record these and their circumstances. It is then able to use this information to assess the failures and

make targeted changes to its neural net, including starting with a different set of weights, or re-training on a specific set of inputs. The agent exhibits better behavior while training, and also learns to navigate effectively more quickly.

10. Summary of Dissertation: Darsana Josyula, ''A Unified Theory Of Acting And Agency For A Universal Interfacing Agent'', University of Maryland, December 2005.

With consumer electronics becoming numerous, varied and complex, the idea of a single, shared, general and flexible interfacing agent to interface human users with the multitude of task-oriented systems or devices seems appealing. Such a universal interfacing agent has to understand user instructions and issue commands to control the task-oriented system to which it is connected, in a manner that the given user desires.

Two important issues that such an agent has to deal with are: (i) how to represent and reason about the tasks that a given device can perform and the results that it can produce and (ii) how to represent and reason about when different tasks are to be performed and whether the tasks have been successful. The dissertation explores these issues in detail and provides a solution to deal with them within a contradiction-tolerant and time-sensitive framework called Active logic.

The solution involves explicitly representing the beliefs, desires, intentions, expectations, observations and achievements of the interfacing agent and reasoning based on these attitudes; the dissertation provides a theory (ALFA) that agents can use in order to perform this reasoning. The theory specifies the interactions between beliefs, observations, desires, intentions, expectations and achievements for a universal interfacing agent, while taking into consideration issues associated with concurrent execution of actions as well as perturbation tolerance. The main characteristics of the theory are: representing and reasoning about concurrent actions and results, dealing with interactions of preconditions of actions or results, dynamic reconsideration of intentions and reasoning using

expectations and achievements.

The dissertation also provides an architecture (DIRECTOR) for implementing agents based on the theory. In this architecture, a meta-cognitive process controls the cognitive activities of the agent. The rudimentary results of implementing the architecture to create a natural language based interfacing agent (ALFRED) are also discussed in the dissertation.

This work also discusses how the agent's underlying Active logic knowledge base evolves during reasoning and provides proofs for properties that the knowledge base exhibits, using a meta-theory that specifies how the knowledge base evolves.

PERSONNEL SUPPORTED

D Perlis (PI)
M Anderson (co-PI)
K Hennacy (Post-doc)
W Chong (GRA)
D Josyula (GRA)

PUBLICATIONS

1. Michael L. Anderson and Tim Oates, eds. Metacognition in Computation: Papers from the 2005 AAAI Spring Symposium. (Menlo Park, CA: AAAI Press), 2005.

   http://www.aaai.org/Press/Reports/Symposia/Spring/ss-05-04.php

2. Michael L. Anderson and Don Perlis. The roots of self-awareness. Phenomenology and the Cognitive Sciences 4(3), 2005.

http://www.agcognition.org/papers/SSC.pdf

3. Michael L. Anderson, Walid Gomaa, John Grant and Don Perlis. On the reasoning of real-world agents: Toward a semantics for active logic. Proceedings of the 7th Annual Symposium on the Logical Formalization of Commonsense Reasoning, Dresden University Technical Report (ISSN 1430-211X), 2005.

   http://www.agcognition.org/papers/CS2005.pdf

4. Michael L. Anderson and Donald R. Perlis. Logic, self-awareness and self-improvement: The metacognitive loop and the problem of brittleness. Journal of Logic and Computation, 15(1), 2005.

   http://www.agcognition.org/papers/04-33-anderson-perlis.pdf

5. Darsana P. Josyula, Michael L. Anderson, and Don Perlis. Designing a Universal Interfacing Agent. In: Proceedings of the Second Language and Technology Conference (L&T'05) on Human Language Technologies as a Challenge for Computer Science and Linguistics, 2005.

   http://www.cs.umd.edu/~darsana/papers/ltc05.pdf

6. Darsana P. Josyula, Michael L. Anderson, and Don Perlis. Metacognition for Dropping and Reconsidering Intentions. In: Papers from the 2005 AAAI Spring Symposium on Metacognition in Computation, pages 62-67, 2005.

   http://www.cs.umd.edu/~darsana/papers/SS605JosyulaD.pdf

7. Michael L. Anderson, Tim Oates, Waiyian Chong and Don Perlis. The Metacognitive Loop I: Enhancing reinforcement learning with metacognitive monitoring and control for improved perturbation tolerance. Submitted.

   http://www.cs.umd.edu/~anderson/papers/JETAI_final.pdf

8. John Grant, Sarit Kraus, and Don Perlis. A Logic-Based Model of Intention Formation and Action for Multi-Agent Subcontracting, Artificial Intelligence 163(2): 163-201 (2005).

http://www.cs.umd.edu/projects/active/publications/papers/finalrevised.pdf

9. Don Perlis. Theory and application of self-reference: logic and beyond. To appear as chapter in book, publ Center for Study of Language and Information (CSLI, Stanford).

http://www.cs.umd.edu/active/publications/papers/bigphi.pdf

10. Michael L. Anderson. Specification of a test environment and performance measures for perturbation-tolerant cognitive agents. Proceedings of the AAAI Workshop on Intelligent Agent Architectures, 2004.

http://www.agcognition.org/papers/aaai_metrics_04.pdf

11. Michael L. Anderson. A flexible approach to quantifying various dimensions of environmental complexity. Proceedings of the NIST Workshop on Performance Measures for Intelligent Agents, 2004.

http://www.activelogic.org/publications/papers/flexible.pdf

12. Michael L. Anderson, Darsana Josyula, Khemdut Purang and Don Perlis. Active Logic for more effective human-computer interaction and other commonsense applications. International Joint Conference on Automated Reasoning, Workshop on Empirically Successful First Order Reasoning, 2004.

http://www.agcognition.org/papers/ijcarfinal.pdf

13. Darsana P. Josyula, Michael L. Anderson, and Don Perlis. Domain-Independent Reason-Enhanced Controller for Task-ORiented systems - DIRECTOR. Proceedings of the National Conference on

Artificial Intelligence (AAAI-04), 2004.

http://www.cs.umd.edu/~darsana/papers/aaai404JosyulaD.pdf

14. Michael L. Anderson and Darsana Josyula and Don Perlis. Talking to Computers. Proceedings of the Workshop on Mixed Initiative Intelligent Systems, IJCAI-03, 2003.

    http://www.agcognition.org/papers/miis.pdf

15. Darsana Josyula and Michael L. Anderson and Don Perlis. Towards domain-independent, task-oriented, conversational adequacy. Proceedings of IJCAI-2003 Intelligent Systems Demonstrations, 2003.

    http://www.agcognition.org/papers/alfred.pdf

16. K. Hennacy, N. Swamy, D. Perlis. RGL Study In A Hybrid Real-Time System, Proceedings of IASTED NCI, 2003

    http://www.cs.umd.edu/projects/active/publications/papers/NCIpaper.pdf

17. Michael L. Anderson and Tim Oates. Pre-linguistic agents will form only ego-centric predicates. Behavioral and Brain Sciences 23(6), 2003.

    http://www.ling.ed.ac.uk/~jim/BBSNEURO/anderson.html

18. Michael L. Anderson. Evans' Varieties of Reference and the anchoring problem. Robotics and Autonomous Systems 43(1-2): 189-92, 2003.

    http://www.cs.umd.edu/~anderson/papers/Evans.pdf

19. Michael L. Anderson. Embodied Cognition: A Field Guide. Artificial Intelligence, 149(1): 91-130, 2003.

http://dx.doi.org/10.1016/S0004-3702(03)00054-7

20. Michael L. Anderson. Representations, Symbols and Embodiment. Artificial Intelligence. 149(1):151-6, 2003.

http://dx.doi.org/10.1016/S0004-3702(03)00056-0

INTERACTIONS/TRANSITIONS

1. Invited talk, Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA. March 2006. Active Logic for Agents. Darsana P. Josyula.
2. Invited talk, BBN Technologies, Cambridge, MA. May, 2005. The metacognitive loop and the problem of brittleness. Michael L. Anderson.
3. Invited talk. Lehigh University. The multiple-modes theory of epistemic access. March 2005. Michael L. Anderson.
4. Talks given at 2005 conferences for the Society for Discourse, Text and Cognition; Interactivist Institute; and McDonnell Project. Michael L. Anderson.
5. Served as chair (M. Anderson) and on program committee (D. Perlis) of the AAAI 2005 Spring Symposium on Metacognition in Computation.
6. Invited talk, DARPA workshop on self-aware systems. Systems that fix themselves. April 2004. Michael L. Anderson.
7. Invited talk, NIST workshop on Performance Measures for Intelligent Agents. Specification of a test environment and performance metrics for perturbation-tolerant cognitive agents. May 2004. Michael L. Anderson.
8. Talks given at the 2004 conferences for AAAI, the Society for Text & Discourse, and the Cognitive Science Society. Michael L. Anderson.
9. Honda Research Institute, USA: initial stage of application of our work on commonsense reasoning, including inconsistencies, for a robot in an indoor environment. Ken Hennacy and Don Perlis.
10. Invited talk, CS and Electrical Computer Engineering Colloquium University of Wyoming, August 2004. Don Perlis
11. Invited talk, Institute for Cognitive Science Colloquium University of Colorado, August 2004. Don Perlis

Patents:

1. System for the automated detection of metalanguage. US60/807,744 (provisional application filed).

Inventions: none

Awards:

1. M. Anderson, recognized as an ''emerging leader under forty'' by the Renaissance Institute (sponsor of the Renaissance Weekend), 2005.

2. M. Anderson, invited participant, McDonnell Project in Philosophy and the Neurosciences workshop for early career researchers, 2005.